STAR DAQ Upgrade R&D Proposal

FY2003 to FY2005

Tonko A. Ljubicic (tonko@bnl.gov)
Brookhaven National Laboratory
September, 10. 2002
Version 1.2

Statement of Value and Justification of the Need

At the STAR Future Upgrades Meeting, June 2002, Bar Harbor, the STAR collaboration expressed a requirement stating DAQ data rates with the TPC on the *order of 1 kHz*.

The current STAR DAQ system is a hierarchically organized system of VME crates connected with a fast Myrinet network. Each crate typically corresponds to a detector or a sub-detector (i.e. a TPC sector) and houses detector specific Receiver Boards. Each Receiver Board holds 3 daughter boards each with an Intel I960HD CPU and 6 STAR-DAQ ASICs to facilitate data formatting as well as preprocessing for the Level III trigger (*cluster finding*). The Receiver Boards for the TPC, SVT, FTPC & SSD are all the same and were custom designed and manufactured.

This system has several bottlenecks due to the original requirements of storing one event per second but being able to analyze 100 events per second into a Level III Trigger farm. The major bottlenecks are: the slow CPU on the daughter boards (66 MHz), the fixed throughput of 100 Hz at the fiber optical input to the Receiver Boards and finally the relatively low bandwidth of the standard VME64 crates (about 40 MB/s). Other bottlenecks are in the architecture of the Event Builder as well as in the Level III processing cluster both of which need to be resized in a scalable manner.

Current DAQ/detector rates are limited to about 50 Hz in the DAQ domain and 100 Hz in the TPC domain. These are *hard* limits and they *cannot be increased* without a *simultaneous* redesign of *both* the DAQ system and the TPC front-end electronics. An incremental upgrade is not possible due to many constraints internal to the current system such as custom ASIC limits, CPU speed limits, network limits etc. A new design with new electronics, computing power and faster network/bus interconnects is necessary.

Increasing the DAQ rates by a *factor of 10 to 20* is a significant challenge and thus necessitates a considerable and well focused Research & Development program.

The main prongs of this R&D would concentrate on

a) STAR-specific data compression computing technology
b) fast interconnect(s) between subparts of the DAQ system
c) the computing technology used for the Level III trigger decision based upon the tracking detector's (assumed to be TPC but need not be) data.
d) computer technology enabling high-rate data aggregation ("Event Building") for further storage

This proposal assumes the STAR TPC as the largest source of data as well as the tracking detector for Level III algorithms. We however plan to design the DAQ upgrade to be independent of the data source as much as possible. Such a design strategy facilitates an eventual change in the basic detector technology in the years to come. Additionally, the developed upgrades will take into account the other current STAR detectors such as SVT, FTPC, EMC etc.

Data Compression Computing Technology

The current STAR DAQ contains about 450 microprocessors managing just the TPC data alone. The most important task these CPUs have is the two-dimensional cluster finding as the important step towards a large data reduction scheme. The current CPUs are Intel I960HD running at 66 MHz and they can handle about 50 Hz of central Au-Au collisions. Increasing this rate by 20 times (to 1 kHz) requires a search for a new processing element which in our view can either be a much faster CPU, a DSP, an FPGA or some combination of the above. Evaluating these technologies is the first step in the process of speeding up the whole system.

Fast Interconnects

The current STAR DAQ system containing the industry-standard VME crates as well as 1 GHz optical connections to the TPC front-end is capable of moving about 100 Hz worth of TPC data. To increase this volume by a factor of 10 a new interconnect technology is necessary, taking advantage of new technology developments on the ten years since the original STAR design. The R&D resources have to be spent evaluating the current cutting-edge networks as well as the board-to-board and chip-to-chip communication paths.

Level III Computing

STAR's Level III computing cluster currently contains about 50 commercial workstations which can handle about 50 Hz central Au+Au collisions. Simply increasing the number of computers to ~1000 to meet the 20 times increase may not be desirable due to per-event latency considerations.

It is necessary to re-evaluate the CPU technology as well as the network interconnecting the future nodes to the rest of the DAQ processing chain. Additional development of the fast tracking software will also be required to better match the required physics performance with the underlying hardware.

In any case, the number of nodes will need to be considerably larger than the current cluster size, and new commercial techniques for packaging, cooling, and management need to be evaluated on a cluster of moderate size before a final design can be seriously specified.

High-rate Event Building

The Event Building is the last stage of the STAR DAQ system and currently can handle about 50 Hz of central Au+Au collisions deploying a large computer workstation with a significant amount of disk and memory. Such a single-station topology can't scale a factor of 20 even assuming future CPU, memory and storage improvements. The only currently known cost-effective approach to this scaling problem is the use of a cluster of high performance workstations working in event-parallel mode. We propose to build a small scale system and evaluate its performance and scalability with current technologies which we can then scale both in performance and price as the construction phase of the project commences.

Since all of these parts would work together as a whole system and are tied together with common interconnects we propose one R&D project to handle these tasks at the same time.

R&D Stages

The R&D effort is assumed staged into three phases matching the Fiscal Years 2003 through 2005.

1) FY2003 Technology Survey

In the first year of this project we plan to evaluate technologies for all the above components using commercially available products. This would mean survey of existing technologies, purchase of vendor-specific development products and platforms, performance evaluation and measurement, code adaptation to differing environments (i.e. implementing the cluster-finding software in FPGAs). For the Level III and Event Builder components we would purchase and evaluate the fastest computers on the market deemed capable and cost-effective for the final design.

A part of the effort would be spent integrating possible technological choices with the new TPC front-end.

2) FY2004 Subprototypes

The second phase would be spent in the design and test of small custom-made units which mock-up the full DAQ chain. I.e. we expect to have a well designed TPC front-end interface by that time and we would spend the effort attempting to evaluate a single vertical slice of the full parallel readout chain.

We expect to have a good understanding of the bottlenecks of the proposed system as well as the Level III farm and the Event Building strategy.

3) FY2005 Final Prototype

The last year will be spent designing and manufacturing the full prototype of a concrete DAQ chain including all the specific interconnects and computing elements. We expect to fully test and debug such a chain as well as measure its performance. By the end of this stage we expect to be prepared for the construction phase of the project where we would manufacture and replicate the processing units based on this final prototype.

Budget Estimate

Category	Comments	FY2003 k$	FY2004 k$	FY2005 k$
Salaries	Engineers and technical support only! Assumes 120 k$/engineer; 90 k$/technician.	90 (3/4 engineer)	330 (2 engineers,1 tech)	330 (2 engineers, 1 tech)
Travel		20	10	10
Purchases		100	300	400
Subcontracts	PCBs…	10	100	100
Overhead	Assumed 35%	80	260	300
Total		***300***	***1000***	***1140***

People

Jeff Landgraf	(BNL)	
Micheal LeVine	(BNL)	
Tonko A. Ljubicic	(BNL)	*Principal Investigator*
David Lynn	(BNL)	
Danny Padrazo?	(BNL)	
Others???		